A SURVEY OF BRANCHING PROCESSES.

JOHN B. SHEWMAKER

A SURVEY OF BRANCHING PROCESSES

* * * * *

John B. Shewmaker

A SURVEY OF BRANCHING PROCESSES

by

John B. Shewmaker

Lieutenant, United States Navy

Submitted in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE

United States Naval Postgraduate School
Monterey, California

1963

A SURVEY OF BRANCHING PROCESSES

by

John B. Shewmaker

This work is accepted as fulfilling
the thesis requirements for the degree of
MASTER OF SCIENCE

from the
United States Naval Postgraduate School

ABSTRACT

A mathematical model of a branching stochastic process utilizing generating functions is presented. The probability distribution of the number of members of the process at discrete time periods, z_n, and the probability of extinction is discussed. When there is a non-zero probability of surviving indefinitely, the normed random variables, $z_n/E(z_n)$, converge with probability one; the cumulative distribution of this random variable is absolutely continuous. The time until extinction and the total number of members of the process is examined when the probability of extinction is one. The distribution of the z_n given that z_n is not zero is discussed for this case. The maximum likelihood estimates for the probabilities involved in the process are determined. An example is given of a branching process in which the probabilities are dependent on time and a solution is found using Laplace transform methods.

TABLE OF CONTENTS

	Page
Abstract	iii
Nomenclature	v
Introduction	1
General	4
Mean greater than one	12
Mean not greater than one	24
Estimation	31
Ion fluctuation	34
Appendix	
Bibliography	38

NOMENCLATURE

Symbol	Meaning
$P(A)$	the probability of the event A
$P(A\ B)$	the conditional probability of A given B
$P(A,B)$	the joint probability of A and B
z_r	the number of members of the r^{th} generation
$X_i^{(n)}$	the number of offspring of the i^{th} member of the n^{th} generation
	the expected value of z_1
	the variance of z_1
$f(s)$	generating function of z_1
p_r	probability of r members in the first generation
$f_n(s)$	generating function of z_n
p_{nr}	probability of r members in the n^{th} generation
a	probability of extinction
w_n	a normed random variable $\left(\frac{z_n}{n}\right)$
w	the limit of w_n as n approaches infinity
$G_n(u)$	the cumulative probability function of w_n
$_n(s)$	the characteristic function of w_n
$G(u)$	asymptotic distribution branching from f(s)
$k(s)$	probability generating function
$H(u)$	the asymptotic distribution branching from k(s)
(s)	the characteristic function of H(u)
a_n	probability that $z_n \neq 0$
p^*_{nr}	probability that z_n equals r given that it is not equal to zero

y_n	number of members of the n^{th} generation given that it is not zero
$F_n(s)$	probability generating function of y_n
N	number of generations to extinction
(s)	moment generating function of N
S	total number of individuals in the process
g(s)	probability generating function of S
z_{mr}	number of individuals in the m^{th} generation who have exactly r descendents in the $(m \ \ 1)^{st}$ generation
S_n	total number of individuals in the first n generations

INTRODUCTION

The present paper is an essay on the theory and application of discrete branching stochastic processes. Most of the asymptotic theory was extracted from the work of Harris [10] who derived many of the relationships for this class of stochastic processes. Let us introduce these processes by means of a historical example.

In 1874, one of the first investigations into discrete branching processes was made by the Rev. H. W. Watson. Dr. F. Galton had posed the problem of extinction of family names to the mathematical community in England since many names, prominent in history, were no longer to be found. The problem was posed as follows: Assume a large nation with N adult males, each with a different surname. What proportion of names would be extinct after r generations? In the r^{th} generation, how many people would have the same name? Rev. Watson [8] derived some relationships that enabled him to compute the probabality of extinction and made some sample calculations to demonstrate the difficulty of computing numerical answers. Notable among these relationships was the functional relation of the probability generating functions involved (section III, theorem 1).

Branching processes received some attention thereafter and the theory was applied to problems of gene mutation [7], population growth [12], cosmic radiation [1], ionization fluctuations [18], neutron production in atomic piles [11], and cascade theory [20].

In this paper we will review the theory of simple discrete branching processes and some of its applications. Let us consider the original problem of survival of family names and state the

basic problem of branching processes in that framework. Suppose an immigrant by the name of Door comes to the United States and there is no one else here having that surname. What is the probability that the name survives r generations and how many men will there be having that last name?

There is a probability that Mr. Door will either not marry or not have any male children; there are other probabilities associated with the number of sons he may have. We will denote these probabilities by p_r, i.e., the probability of no sons will be p_0, of r sons, p_r. We need a symbol for the number of individuals with the name of Door in the n^{th} generation, so let z_n be the number of individuals with the name of Door in the n^{th} generation. Then we can write $P(z_1 = r) = p_r$ indicating that the probability of r individuals with the name of Door in the first generation is p_r. If Mr. Door does have sons, each of them may have children and the process is continued. This may be represented as a tree type graph with an initial node representing his sons, and so on. A method of analysis of the problem based on graph theory was proposed by Otter [19]. In this paper the method of generating functions is used.

We will assume that the number of sons born to one of Mr. Door's sons follows the same probability law as that of Mr. Door. In other words, the probability that Door, jr., has n children is p_n. Then, if $X_i^{(n)}$ is the number of offspring of the i^{th} individual in the n^{th} generation, we have $P(z_{n+1} = r \mid z_n = m) = P(\sum_{i=1}^{m} X_i^{(n)} = n)$.

This is the outline of the problem, the notation, and the assumptions made in this paper. Before taking up the mathematical theory, some other problems will be posed in this framework.

In atomic piles, the number of free neutrons is an important factor. If we start with one neutron and a fission takes place, the probability that r neutrons are released is a nuclear constant which must be empirically determined for the amount and arrangement of active material. The probability of no fission (no descendents) is the average probability of leakage or absorption. An atomic pile in a nuclear reactor is subcritical when the reactor is not operating and there is no possibility of an explosion, i.e., the probability that the number of fissions increases sufficiently to cause an explosion is zero; the mean of the probability law for a subcritical mass is less than one. As the amount of fissionable material is increased, the probability of an explosion finally becomes greater than zero. When this occurs, the mass is termed supercritical and the mean of the probability law for this situation is found to be greater than one. A nuclear reactor is termed critical when it is operating; the mean of its probability distribution is exactly one.

The organization of the paper is as follows: The general mathematical character of a discrete branching process is presented in section I. It will be seen that the mean number of progeny per individual plays an important role in the asymptotic theory. This theory is presented in section II for the mean of the probability distribution greater than one and in section III if it is not greater than one. The maximum likelihood estimation of the distribution of progeny per individual is given in section IV. An example of an ionization cascade is presented in section V. This is a more general branching process than those treated in the model used in the earlier sections and its treatment suggests how the methods of Markov processes may be applied.

GENERAL

The structure of branching processes was presented in the introduction where several assumptions were made and some notation was introduced. It is appropriate now to summarize the assumptions made in this paper and to extend the notation.

Consider two related classes of random variables:

z_n the number of "individuals" in the n^{th} generation

$X_i^{(n)}$ the number of "sons" of the i^{th} individual in the n^{th} generation,

They are related by the equations:

$$z_n = \sum_{i=1}^{z_{n-1}} X_i^{(n-1)} \qquad \text{for } n = 1,2,3,\ldots$$

The quantity z_0 is the number of the initial group who start the process; z_0 is assumed to be one, since this does not affect the generality of the theory as is pointed out at the end of this section. The probability distribution of the first generation is given by p_r, i.e., the probability that there are r members of the first generation is p_r.

The following assumptions are made throughout this paper:

(1) Given z_n is equal to m, the random variable, z_{n+1}, is the sum of m independent random variables, each one of which has the same distribution as z_1 : $P(X_i^{(n)} = r) = p_r$

(2) The variance of z_1 is finite.

(3) $p_r \neq 1$, for all r ; otherwise the process degenerates to a simple arithmetic calculation.

(4) $p_0 + p_1 \neq 1$, otherwise we have the binomial case whose treatment in the context of branching processes is trivial.

The following notation is used throughout:

$\mu = E(z_1) = \sum_{r=0}^{\infty} r p_r$, the mean of z_1

$\sigma^2 = Var(z_1) = \sum_{r=0}^{\infty} (r^2 p_r) - \mu^2$ the variance of z_1

$f(s) = \sum_{r=0}^{\infty} p_r s^r$ the probability generating function of z_1 (s is a complex variable).

$$p_{nr} = P(z_n = r)$$

$f_n(s) = \sum_{r=0}^{\infty} p_{nr} s^r$ the probability generating function of z_n.

From the assumptions listed above, it is clear that : $f_o(s) = s$, $p_{1r} = p_r$, $f'(s)$ and $f''(s)$ are continuous on the set of points consisting of $|s| < 1$ and $s=1$.

Several authors (see [2], [9], [11]) have examined the generating functions and have found some basic relationships. Some of these are presented below.

Theorem 1. $f_n(s)$ is the n^{th} functional iterate of $f(s)$:

$$f_n(s) = f(f_{n-1}(s)) = f_{n-1}(f(s))$$

Proof: From the initial assumptions, the $X_i^{(n)}$ are independent for all i and n, $P(X_i^{(n)} = k) = p_k$, $z_n = \sum_{i=1}^{z_{n-1}} X_i^{(n-1)}$

It follows that $P(z_{n+1} = k \mid z_n = j)$ is the coefficient of s^k in the power series expansion of $(f(s))^j$ since the generating function of the sum of j mutually independent random variables is the product of their generating functions. See Feller [6, p. 250ff.].

Hence, $f_2(s) = \sum_{j=0}^{\infty} p_j (f(s))^j = f(f(s))$

$$f_3(s) = \sum_{j=0}^{\infty} p_{2j} [f(s)]^j = f_2[f(s)] = f[f(f(s))] = f[f_2(s)]$$

.

.

.

$$f_n(s) = \sum_{j=0}^{\infty} p_{n-1}[f(s)]^j = f_{n-1}[f(s)] = f_{n-1}(f(\ldots(f(f(s)))\ldots))$$
$$= f(f_{n-1}(s))$$

The coefficient of s^r in $f_n(s)$ is the probability that there are r individuals in the n^{th} generation. Thus the questions posed by Dr. Galton may be answered in the framework of the present model. The constant term, p_{no}, is the probability that the family name is extinct by the n^{th} generation.

The mean and the variance of z_n can be determined using theorem 1.

Theorem 2. $E(z_n) = \mu^n \qquad Var(z_n) = \dfrac{\sigma^2 \mu^n(\mu^n - 1)}{\mu^2 - \mu} \qquad \mu \neq 1$

$$= n\sigma^2 \qquad \mu = 1$$

Proof: $f_n'(1) = \sum_{r=0}^{\infty} r p_{nr} = E(z_n)$

$$f_n''(1) = \sum r(r-1) p_{nr} = E(z_n^2) - E(z_n)$$

Therefore, $Var(z_n) = f_n''(1) + f_n'(1) - [f_n'(1)]^2$.

By theorem 1 $\quad f_{n+1}(s) = f_n[f(s)] = f[f_n(s)]$

Taking derivatives, $f_{n+1}'(s) = f_n'[f(s)]f'(s) = f'[f_n(s)]f_n'(s)$

$$f''_{n+1}(s) = f''[f(s)][f'(s)]^2 + f'_n[f(s)]f''(s)$$

$$= f''[f_n(s)][f_n'(s)]^2 + f'[f_n(s)]\, f_n''(s)$$

Hence $\quad f_{n+1}(1) = \mu E(Z_n) = \mu^{n+1}$

$$f''_n [f(1)][f'(1)]^2 + f'_n[f(1)]f''(1) = f''[f_n(1)][f'_n(1)]^2 + f'[f_n(1)]f_n''(1)$$

Solving for $f''_n(1)$ gives

$$\therefore f''_n(1) = \frac{f''(1)[f'_n(1)]^2 - f'_n(1)f''(1)}{[f'(1)]^2 - f'(1)}$$

$$= \frac{(\sigma^2 - \mu + \mu^2)(\mu^{2n} - \mu^n)}{\mu^2 - \mu}$$

$$= \frac{\sigma^2 \mu^n(\mu^n - 1)}{\mu^2 - \mu} - \mu^n + \mu^{2n}$$

Thus $\mathrm{Var}(z_n) = \dfrac{\sigma^2\mu^n(\mu^n - 1)}{\mu^2 - \mu}$ $\qquad$ for $\mu \neq 1$

For the case $\mu = 1$, L'Hopital's Rule can be used:

$$\left.\frac{\sigma^2(2n\mu^{2n-1} - n\mu^{n-1})}{2\mu - 1}\right|_{\mu=1} = \frac{\sigma^2(2n-n)}{2-1}$$

and $\mathrm{Var}(z_n) = n\sigma^2$ $\qquad$ for $\mu = 1$

Higher moments, if they exist, may be found by a similar process utilizing the higher derivatives of $f_{n+1}(s)$.

A subject of great interest in any branching process is the probability of extinction of the species. The Fundamental Theorem

of branching processes tells us that this probability is a fixed point of the generating function. The proof given follows that of Bharucha-Reid [2]. First, we note that if "a" is the probability of extinction, then $a = \lim_{n\to\infty} p_{no} = \lim_{n\to\infty} f_n(0)$.

<u>Theorem 3.</u> The probability of extinction, a, is the smallest positive number, p, such that :

$$p = f(p) = \sum_{j=0}^{\infty} p_j p^j$$

Further, a = 1 if and only if $\mu \leq 1$.

<u>Proof:</u> We see that a = 0 if and only if $p_o = 0$ and f(1) = 1, so that $0 \leq a \leq 1$ as it should be.

To show that a = f(a), we note that since $p_{no} = f_n(0)$,

$$p_{n+1,0} = f[f_n(0)]$$

so $a = \lim_{n\to\infty} p_{no} = \lim_{n\to\infty} f_n(0) = \lim_{n\to\infty} f(f_{n-1}(0)) = \lim_{n\to\infty} f(p_{n-1,0}) = f(a)$.

Now f(s) is a power series in s with positive coefficients and hence strictly increasing for real arguments. If $0<a<b$, then $f(a) < f(b)$.

Also, since $p_{no} \geq 0$ for all n, then $a \geq 0$.

Let $p \geq 0$ be such that $p = f(p)$. It follows that

$$p_{10} = f(0) \leq f(p) = p$$

Assuming $p_{no} \leq p$, then $f_n(0) \leq p$ and

$$p_{n+1,0} = f_{n+1}(0) = f[f_n(0)] \leq f(p) = p$$

So, by induction, $p_{no} \leq p$ for all n.

Then $\lim_{n\to\infty} p_{no} \leq p$ and $a \leq p$.

It remains to show that a = 1 if and only if $\mu \leq 1$.

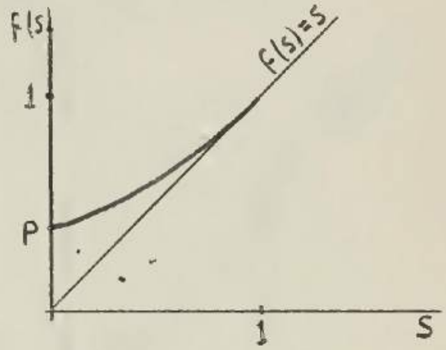
Assume first that a = 1 and show that $\mu \leq 1$

Now, it may be noted that:

$f'(s) \geq 0$, $f''(s) \geq 0$, so $f(s)$ is convex in $(0,1)$

$f(0) = p_o > 0$, $f(1) = 1$

$f(s) > s$ for s in $(0,1)$ since 1 isthe minimum positive number satisfying $s = f(s)$.



Therefore, $1 - f(s) < 1-s$ and.

$\frac{1-f(s)}{1-s}$ is bounded by 1 and is monotonically increasing with s for s in $(0,1)$.

Therefore, $f'(1)$ exists, $f'(1) \leq 1$ and $\mu \leq 1$.

Now assume $\mu \leq 1$. Then, since $f(s)$ is convex, $f'(s)$ is either constant in $(0,1)$ or strictly increasing with s. By assumption, $p_1 \neq 1$ so $p_o > 0$ and $f'(s) < 1$ in either case for s in $(0,1)$.

By the mean value theorem, $f(1) - f(s) = f'(x)(1-s)$ where $s \leq x \leq 1$

Then, $1 - f(s) = f'(x)(1-s)$ and

$$1 - f(s) < 1 - s$$

$$f(s) > s \quad \text{for} \quad 0 \leq s < 1$$

Therefore, a = 1.

Using theorem 3, it is a simple matter to determine the limits of the p_{nr} for $r \neq 0$.

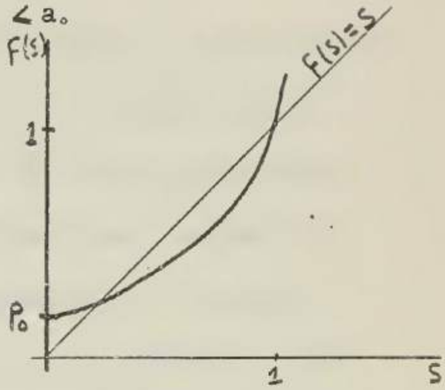
<u>Theorem 4.</u> A discrete branching process either dies out or becomes extremely large, i.e., $\lim_{n\to\infty} p_{nr} = 0$ for $r \neq 0$.

Proof: Let t be a real number such that $0 \leq t < 1$ and recall the iteration property of $f_n(t)$.

Case I. $\mu \leq 1$ Then $f(t) > t$ (see proof of theorem 3), and $f(t) < f_2(t) < \ldots < f_n(t) < \ldots < 1$.

Therefore, $\lim_{n\to\infty} f_n(t) = 1 = a$.

Case II. $\mu > 1$ and $t < a$ $(p_o > 0)$. Then $f(t) > t$ since $f(s)$ is convex and $f(t) < f_2(t) < \ldots < f_n(t) < \ldots < a$.



Therefore, $\lim_{n\to\infty} f_n(t) = a$.

Case III. $\mu > 1$ and $t > a$.

Then $f(t) < t$ due to the convexity of $f(s)$.

Hence $f(t) > f_2(t) > \ldots > f_n(t) > \ldots > a$.

Therefore, $\lim_{n\to\infty} f_n(t) = a$.

Now $f_n(t) = \sum_{r=0}^{\infty} p_{nr} t^r = p_{no} + \sum_{r=1}^{\infty} p_{nr} t^r$ by definition. Taking the limit of the last expression gives:

$$\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty} p_{no} + \lim_{n\to\infty} \sum_{r=1}^{\infty} p_{nr} t^r = a$$ from the above three cases.

Therefore, $\lim_{n\to\infty} \sum_{r=1}^{\infty} p_{nr} t^r = 0$. Since the series converges uniformly in an interval one may pass to the limit before summing.

Therefore $\sum_{r=1}^{\infty} \lim_{n\to\infty} p_{nr} t^r = 0$ and $\lim_{n\to\infty} p_{nr} = 0$ for $r \neq 0$ by the uniqueness of power series.

Theorems 3 and 4 give us some interesting information. If we are working with an atomic pile, the probability is one that it will die out unless it is supercritical; then the probability that it will blow up is 1 - a.

Lotka [17] found that the probability law of the number of sons a man might have was closely approximated by a type of geometric distribution, namely, $p_o = 0.4825$, $p_k = (0.2126)(0.5893)^{k-1}$ $k \geq 1$ based on a 1920 census. It follows that:

$$f(s) = 0.4825 + \frac{0.2126s}{1-0.5893s} \qquad \mu = f'(s)\Big|_{s=1} = 1.2622$$

Solving $f(s) = s$, it is found that $a = 0.6226$. Based on these values the probability of extinction of a family name is $(0.6226)^n$ where n is the number of male members of the family at the present time. This formula follows from the fact that if there are r members of the zeroth generation then the generating function for the first generation is $(f(s))^r$. Then $f_n(s) = f_{n-1}((f(s))^r)$; if $f(a) = a$, then $(f(a))^r = a^r$ and a^r is the probability of extinction.

MEAN GREATER THAN ONE

In this section we will consider the case where μ is strictly greater than one and the family has a non-zero probability of surviving indefinitely. Since we have determined that z_n may become extremely large, we will work with a normed random variable.

Let $w_n = \dfrac{z_n}{n}$. The mean and variance of w_n are easily determined.

$$E(w_n) = E\left(\frac{z_n}{\mu^n}\right) = \frac{1}{\mu^n} E(z_n) = 1$$

$$E(w_n^2) = E\left(\frac{z_n^2}{\mu^{2n}}\right) = \frac{1}{\mu^{2n}}\left[\frac{\sigma^2\mu^n(\mu^n-1)}{\mu^2-\mu} + \mu^{2n}\right]$$

$$= \frac{\sigma^2}{\mu^2-\mu}\left(1-\frac{1}{\mu^n}\right) + 1$$

$$\mathrm{Var}(w_n) = \frac{\sigma^2}{\mu(\mu-1)}\left(1-\frac{1}{\mu^n}\right)$$

Treatment of the convergence of w_n may be found in Harris [10] and Bharucha-Reid [2]. We will prove first, the convergence in probability and then the convergence with probability one.

<u>Theorem 5</u>. If $\mu > 1$, the random variables, w_n, converge in probability to a random variable w.

<u>Proof</u>. Let n, m, be integers and $n > m$.

$$E(z_n | z_m = r) = \sum_{j=1}^{\infty} j\,[\text{coefficient of } s^j \text{ in}(f_{n-m}(s))^r]$$

$$= \frac{d}{ds}\left(f_{n-m}(s)\right)^r\Big|_{s=1}.$$

$$E(z_n|z_m=r) = r[f_{n-m}(s)]^{r-1} f'_{n-m}(s)\Big|_{s=1}$$

$$= r(1)^{r-1}\mu^{n-m}$$

$$= r\mu^{n-m}$$

$$E(z_n z_m) = \sum_{r=0}^{\infty} P_{mr} E(r z_n|z_m=r)$$

$$= \sum_r P_{mr} r^2 \mu^{n-m}$$

$$= \mu^{n-m} E(z_m^2)$$

$$E(W_n W_m) = \frac{1}{\mu^{n+m}} E(Z_n Z_m)$$

$$= \frac{1}{\mu^{n+m}} \mu^{n-m} E(Z_m^2)$$

$$= E(W_m^2)$$

$$E[(W_n - W_m)^2] = E(W_n^2) - E(W_m^2)$$

$$= \frac{\sigma^2}{\mu^2-\mu}\left(\frac{1}{\mu^m} - \frac{1}{\mu^n}\right)$$

$$\lim_{m,n\to\infty} E[(W_n - W_m)^2] = 0$$

Hence the random variables converge in mean square and by a theorem of Kolmogorov [13, p.34, I] , in probability.

Convergence with probability one may be proven with the aid of Doob's [4] Martingale convergence theorem.

A discrete parameter martingale is a stochastic process such that:

a) $E(|y_n|) < \infty$ for all n

b) $E(y_{n+1} | y_n, y_{n-1}, \ldots, y_1) = y_n$

The theorem may be stated as follows. Let $(y_n, n \geq 0)$ be a discrete parameter martingale. Then $E(|y_0|) \leq E(|y_1|) \leq E(y_2) \leq \ldots$

If $\lim_{n\to\infty} E(|y_n|) = k < \infty$, then $\lim_{n\to\infty} y_n = y^*$ exixts with probability one and $E(|y^*|) \leq k$.

Considering our random variable w_n, $E(w_n) = 1$ for all n,

$$E(W_{n+1} | W_n) = E\left(\frac{Z_{n+1}}{\mu^{n+1}} \Big| \frac{Z_n}{\mu^n}\right)$$

$$= \frac{Z_n}{\mu^n}$$

$$= W_n$$

Hence the (w_n) form a discrete parameter martingale and we have

Theorem 4a. If $\mu > 1$, the random variables, w_n, converge with probability one to a random variable w.

Since w_n converges to w, we may discuss the probability distribution of w.

Let $G_n(u) = P(w_n \leq u)$, $\varphi_n(s) = E(e^{w_n s}) = \int_0^\infty e^{su}\, dG_n(u)$, $G(u) = P(w \leq u)$, $\varphi(s) = \int_0^\infty e^{su}\, dG(u)$. The distribution G(u) can be called the asymptotic distribution branching from f(s). There is a very interesting functional relation between the moment generating function $\varphi(s)$ of w and the probability generating function f(s) as given by the following theorem.

Theorem 5. $\varphi(\mu s) = f(\varphi(s))$

Proof: $\varphi_n(s) = E(e^{w_n s}) = E(e^{\frac{z_n s}{\mu^n}}) = \sum_{j=0}^{\infty} p_{nj} \exp[j \frac{s}{\mu^n}]$

$$= f_n(e^{\frac{s}{\mu^n}})$$

$$\varphi_{n+1}(\mu s) = f_{n+1}(e^{\frac{s}{\mu^n}}) = f[f_n(e^{s/\mu^n})]$$

$$= f[\varphi_n(s)]$$

$$\lim_{n\to\infty} G_n(s) = G(s) \quad \text{,therefore,}$$

$$\lim_{n\to\infty} \varphi_n(s) = \lim_{n\to\infty} \int_0^{\infty} e^{su} dG_n(u)$$

$$= \int_0^{\infty} e^{su} dG(u) = \varphi(s)$$

Upon taking the limit of $\varphi_{n+1}(\mu s) = f[\varphi_n(s)]$ the desired result is obtained.

It also follows that the k^{th} moment of w exists and is finite if and only if the k^{th} moment of z_1 exists and is finite.

Since $G_n(0) = P(w_n \leq 0) = p_{no}$, then $\lim_{u\to 0} G_n(0) = a$ and $P(w = 0) = a$. The following theorem will be used later.

Theorem 6. Let $G_1(u)$, $G_2(u)$ be distributions having equal first moments and finite second moments such that their characteristic functions $\varphi_1(it)$ and $\varphi_2(it)$ satisfy $\varphi_r(it\mu) = f[\varphi_r(it)]$ r = 1,2

Then $G_1(u) = G_2(u)$

<u>Proof</u>: $\varphi(it) = 1 + \sum_{\nu=1}^{K-1} \frac{\alpha_\nu}{\nu!}(it)^\nu + o(t^K)$

for $|t|$ small by a special form of McLaurin's theorem for small values of $|t|$. See [Cramer 3, p.27]. Then,

$$|\varphi_1(it) - \varphi_2(it)| = \left| 1 + \sum_{n=1}^{2} \frac{\alpha_n^{(1)}}{n!}(it)^n + o(t^3) - 1 \right.$$
$$\left. - \sum_{n=1}^{2} \frac{\alpha_n^{(2)}}{n!}(it)^n + o(t^3) \right| \quad \text{where } \varphi_j^{(n)}(it) = i^n \alpha_n^{(j)}$$

$$|\varphi_1(it) - \varphi_2(it)| = \left| (\alpha_1^{(1)} - \alpha_1^{(2)}) it - \tfrac{1}{2}(\alpha_2^{(1)} - \alpha_2^{(2)}) t^2 + o(t^3) \right|$$
$$= t^2 \left| \tfrac{1}{2}(\alpha_2^{(2)} - \alpha_1^{(2)}) + o(t) \right|$$

Therefore, $|\varphi_1(it) - \varphi_2(it)| = t^2 \beta(t)$ where $\lim_{t\to 0} \beta(t) < \infty$

Then $|\varphi_1(it\mu) - \varphi_2(\mu it)| = |F[\varphi_1(it)] - F[\varphi_2(it)]|$ by theorem 5

$$\leq \mu |\varphi_1(it) - \varphi_2(it)| \quad \text{since } |F'(s)| \leq \mu \text{ when } |s| \leq 1$$

Hence $|t^2 \mu^2 \beta(\mu t)| \leq \mu |t^2 \beta(t)|$

and $\mu|\beta(\mu t)| \leq |\beta(t)|$ implying that $\mu|\beta(0)| \leq |\beta(0)|$

But $\mu > 1$ so $\beta(t)$ must be identically zero.

Therefore $|\varphi_1(it) - \varphi_2(it)| = 0$

and $\varphi_1(it) \equiv \varphi_2(it)$

Finally $G_1(u) \equiv G_2(u)$ by the uniqueness of moment generating functions.

In order to determine some properties of $G(u)$, it will be convenient to work with a closely related function $H(u)$ which is defined as the asymptotic distribution branching from $k(s)$ where

$$k(s) = \frac{F[s(1-a)+a] - a}{1-a}$$

To demonstrate that k(s) is a generating function use the fact that $\mu > 1$ implies $f(a) = a < 1$. Then

$$k(0) = \frac{f(a) - a}{1-a} = 0$$

$$k(1) = \frac{f(1) - a}{1-a} = 1$$

Clearly k(s) is monotonically increasing since, if $t < s$, then

$$k(t) = \frac{f[t(1-a)+a]-a}{1-a} \leq \frac{f[s(1-a)+a]-a}{1-a} = k(s)$$

Therefore k(s) is a probability generating function.

It is interesting to note the following properties.

$$k'(s) = \frac{f'[s(1-a)+a](1-a)}{1-a} = f'[s(1-a)+a]$$

$$k'(1) = f'(1) = \mu$$

$$k''(s) = f''[s(1-a)+a](1-a)$$

$$k''(1) = (\sigma^2 + \mu^2 - \mu)(1-a)$$

$$k''(1) + k'(1) - [k'(1)]^2 = \sigma^2(1-a) - a(\mu^2 - \mu)$$

$$k(s) = \frac{1}{1-a}\left(\sum_{r=0}^{\infty} p_r (s(1-a)+a)^r\right) = \sum_{r=0}^{\infty} q_r s^r$$

The coefficient of s^k is $(1-a)^{k-1}\left(p_k + p_{k+1}ka + \ldots + \binom{k+n}{k} p_{k+n} a^n + \ldots\right)$ and $q_1 = p_1 + 2p_2 a + 3p_3 a^2 + \ldots = f'(a)$.

Let $\psi(s) = \int_0^{\infty} e^{us} dH(u)$. By a line of argument similar to the proof of theorem 5, it can be shown that $\psi(\mu s) = k[\psi(s)]$.

To demonstrate that k(s) is a generating function use the fact that $\mu > 1$ implies $f(a) = a < 1$. Then

$$k(0) = \frac{f(a) - a}{1-a} = 0$$

$$k(1) = \frac{f(1) - a}{1-a} = 1$$

Clearly k(s) is monotonically increasing since, if $t < s$, then

$$k(t) = \frac{f[t(1-a)+a]-a}{1-a} \leq \frac{f[s(1-a)+a]-a}{1-a} = k(s)$$

Therefore k(s) is a probability generating function.

It is interesting to note the following properties.

$$k'(s) = \frac{f'[s(1-a)+a](1-a)}{1-a} = f'[s(1-a)+a]$$

$$k'(1) = f'(1) = \mu$$

$$k''(s) = f''[s(1-a)+a](1-a)$$

$$k''(1) = (\sigma^2 + \mu^2 - \mu)(1-a)$$

$$k''(1) + k'(1) - [k'(1)]^2 = \sigma^2(1-a) - a(\mu^2 - \mu)$$

$$k(s) = \frac{1}{1-a}\left(\sum_{r=0}^{\infty} p_r (s(1-a)+a)^r\right) = \sum_{r=0}^{\infty} q_r s^r$$

The coefficient of s^k is $(1-a)^{k-1}\left(p_k + p_{k+1}ka + \ldots + \binom{k+n}{k} p_{k+n} a^n + \ldots\right)$ and $q_1 = p_1 + 2p_2 a + 3p_3 a^2 + \ldots = f'(a)$.

Let $\psi(s) = \int_0^\infty e^{vs} dH(v)$. By a line of argument similar to the proof of theorem 5, it can be shown that $\psi(\mu s) = k[\psi(s)]$.

We now proceed to determine a representation for the random variable w. First we determine the relationship between $\psi(s)$ and $\varphi(s)$.

Theorem 7. $$\psi(s) = \frac{\varphi[(1-a)s] - a}{1-a}$$

Proof: Let $\psi_1(s) = \dfrac{\varphi[(1-a)s] - a}{1-a}$

Then $$\psi_1'(s) = \frac{\varphi'[(1-a)s](1-a)}{1-a} = \varphi'[(1-a)s]$$

and $\psi_1'(0) = \varphi'(0) = \mu$. Therefore ψ_1 and ψ have equal first moments. Since $\psi_1''(s)\Big|_{s=0} = \varphi''[(1-a)s](1-a)\Big|_{s=0} < \infty$

ψ_1 and ψ have finite second moments.

$$\psi_1(\mu s) = \frac{\varphi[(1-a)s\mu] - a}{1-a} = \frac{f[\varphi[(1-a)s]] - a}{1-a}$$

$$= \frac{f[(1-a)\psi_1(s)+a] - a}{1-a}$$

$$\psi_1(\mu s) = K[\psi_1(s)]$$

Therefore ψ and ψ satisfy the same functional relationship with k(s). Therefore by theorem 6, $\psi(s) = \psi_1(s)$.

Now we can determine the relationship between H(u) and G(u).

Theorem 8. $$H(u) = \frac{G\left(\frac{u}{1-a}\right) - a}{1-a}.$$

Proof: From theorem 7, $\psi(s) = \dfrac{\varphi[(1-a)s]-a}{1-a}$.

By definition of generating functions $\psi(s) = \int_0^\infty e^{su}\, dH(u)$

and $$\frac{1}{1-a}\left[\int_0^\infty e^{u(1-a)s}\, dG(u) - a\right] = \frac{\varphi[(1-a)s]-a}{1-a}$$

Hence $$\int_0^\infty e^{su}\, dH(u) = \frac{1}{1-a}\left[\int_0^\infty e^{u(1-a)s}\, dG(u) - a\right]$$

$$\int_0^\infty (1-a)e^{su}\, dH(u) = \int_0^\infty e^{u(1-a)s}\, dG(u) - a$$

Let $\varepsilon(u)$ be defined as follows: $\varepsilon(u) = 0$ for $u < 0$

$\varepsilon(u) = a$ for $u > 0$

Then $a = \int_{0^-}^\infty d\,\varepsilon(u) = \int_{0^-}^\infty e^{su}\, d\varepsilon(u)$.

Therefore a is the Laplace transform of $\varepsilon(u)$.

Hence $\int_0^\infty e^{su}(1-a)\,dH(u) = \int_0^\infty e^{su}\, dG\left(\frac{u}{1-a}\right) - \int_{0^-}^\infty e^{su}\, d\varepsilon(u)$

and $\int_{0^-}^\infty e^{su}\, d[(1-a)H(u)] = \int_{0^-}^\infty e^{su}\, d\left[G\left(\frac{u}{1-a}\right) - \varepsilon(u)\right]$

Therefore $(1-a)H(u) = G\left(\frac{u}{1-a}\right) - \varepsilon(u)$

and $H(u) = \frac{1}{1-a}\left[G\left(\frac{u}{1-a}\right) - a\right]$ for $u > 0$

Clearly H(u) = 0 and H(∞) = 1 as it should.

To put the relationship of H(u) and G(u) in terms of random variables, we prove the following theorem.

Theorem 9. The random variable w is distributed as the product of two indepeident random variables, w_0 and w', where $P(w_0 = 0) = a$, $P(w_0 = \frac{1}{1-a}) = 1-a$ and w' has the distribution H(u).

Proof: $E(e^{w_0 w' s}) = a + (1-a)\int_0^\infty e^{\frac{su}{1-a}}\, dH(u)$

$= a + (1-a)\int_0^\infty e^{sv}\, dH(v(1-a))$

$= \int_0^\infty e^{su}\, dG(u)$

$= \varphi(s)$

So, $P(w(1-a) \leq u \mid w > 0) = H(u)$.

For later use a theorem due to Harris [10] is given with some remarks. Let $\gamma = \log_{\mu}\left(\frac{1}{q_1}\right) = \log_{\mu}\left(\frac{1}{f'(a)}\right)$. If $q_1 = 0$, let $\gamma = \infty$.

Theorem 10. If $\gamma < \infty$, $\mathrm{Re}(s) \leq 0$, $s \neq 0$, then $\psi(s) = \dfrac{M(s)}{|s|^{\gamma}} + M_0(s)$

where $M(s)$ is continuous for $s \neq 0$, $M(\mu s) = M(s)$; $M_0(s) = O\left(\frac{1}{|s|^{2\gamma}}\right)$, $|s| \to \infty$

$M(-s) = M(s)$.

Remarks: (a) Under the conditions of the theorem $M(s)$ is real and poaitive when s is real and negative.

(b) If $E|Z_1^r| < \infty$, the r^{th} derivative of $\psi(s)$ satisfies

$$\left|\psi^{(r)}(s)\right| = O\left(\frac{1}{|s|^{r+\gamma}}\right) \qquad |s| \to \infty$$

(c) If $\gamma = \infty$, $\psi(s)$ and as many derivatives as exist approach zero exponentially as $|s| \to \infty$.

We will now determine some of the properties of $G(u)$; however, it will be convenient to prove theorems for $H(u)$ and then interpret the results in terms of $G(u)$. Let $h(u) = H'(u)$.

Theorem 11. $H(u)$ is absolutely continuous.

Proof: Cramer [3] breaks up an arbitrary cumulative distribution function $F(x)$ into three pieces:

$$F(x) = a_1 F_1(x) + a_2 F_2(x) + a_3 F_3(x)$$

where $a_1 + a_2 + a_3 = 1 \qquad a_i \geq 0, \quad i = 1, 2, 3$

$F_1(x)$ is absolutely continuous.

$F_2(x)$ is a step-function and is equal to the sum of the saltuses of $F(x)$ at all discontinuities which are less than or equal to x.

$F_3(x)$, the "singular" component, is a continuous function having, almost everywhere, a derivative equal to zero.

$a_2 \varphi_2(t) = \sum_{n=0}^{\infty} p_n e^{\iota t x_n}$ is the sum of an absosutely convergent trigonometric series and is thus an almost periodic function which comes as close to a_2 as we please for arbitrarily large values of t. By theorem 10, $\lim_{t\to\infty} \psi(t) = 0$. Therefore, for H(u), $a_2 = 0$ and H(u) is continuous.

Let $h_m(v) = \frac{1}{2\pi} \int_{-m}^{m} e^{-\iota t v} \psi(\iota t)\, dt \qquad m = 1, 2, \ldots$

Let $dv = e^{-\iota t v} dt \qquad u = \psi(\iota t)$

$v = -\frac{1}{\iota v} e^{-\iota t v} \qquad du = d[\psi(\iota t)]$

$$h_m(v) = -\frac{1}{2\pi\iota v}\left[e^{-\iota t v}\psi(\iota t)\right]_{-m}^{m} + \int_{-m}^{m} \frac{1}{2\pi\iota v} e^{-\iota t v} \frac{d\psi(\iota t)}{dt}\, dt$$

$$\text{(A)} \quad h_m(v) = -\frac{1}{2\pi\iota v}\left[\psi(\iota m)e^{-\iota m v} - \psi(-\iota m)e^{\iota m v}\right] + \int_{-m}^{m} \frac{e^{-\iota t v}}{2\pi\iota v} \frac{d\psi(\iota t)}{dt}\, dt$$

Examine each term of (A) as $m \to \infty$

First $\psi(\iota m)e^{-\iota m v} = \left[\frac{M(\iota m)}{|m|^{\gamma}} + M_0(s)\right]e^{-\iota m v}$ by theorem 10.

$$= \left[|m|^{-\gamma} M(\iota m) + O\left(\frac{1}{|m|^{2\gamma}}\right)\right] e^{-\iota m v}$$

As we pass to the limit, this term is a function of continuous functions; hence, it must be continuous. Similarly the second term leads to the expression $\psi(-\iota m)e^{-\iota m v} = \left[|m|^{-\gamma} M(\iota m) + O\left(\frac{1}{|m|^{2\gamma}}\right)\right]e^{\iota m v}$

which converges to a continuous function.

The last term in (A) is $\int_{-m}^{m} \frac{e^{-\iota t v}}{2\pi\iota v} \frac{d\psi(\iota t)}{dt}\, dt$

Since $\left|\int F(x)\,dx\right| \leq \int |F(x)|\,dx$, it follows that

$$\left|\int_{-m}^{m} \frac{e^{-\iota t v}}{2\pi\iota v} \frac{d\psi(\iota t)}{dt}\, dt\right| \leq \int_{-m}^{m} \left|\frac{e^{-\iota t v}}{2\pi\iota v}\right| \left|\frac{d\psi(\iota t)}{dt}\right| dt$$

$$\leq \frac{1}{2\pi\iota v} \int_{-m}^{m} \left|\frac{d\psi(\iota t)}{dt}\right| dt$$

$$\left| \int_{-m}^{m} \frac{e^{-itu}}{2\pi iu} \frac{d\psi(it)}{dt}\, dt \right| \leq k\, m\, O\left(\frac{1}{|m|^{1+\gamma}}\right)$$

by corollary b to theorem 10 for some constant k.

$$\leq k\, O\left(\frac{1}{|m|^{\gamma}}\right)$$

Hence this term approaches zero as $m \to \infty$. Therefore, the continuous functions $h_m(u)$ converge uniformly in some interval, $u_1 \leq u \leq u_2$, to a continuous function $h(u)$.

$$H(u_2) - H(u_1) = \lim_{m\to\infty} \int_{-m}^{m} \frac{e^{-itu_2} - e^{-itu_1}}{-2\pi i t}\, \psi(it)\, dt$$

$$= \lim_{m\to\infty} \frac{1}{2\pi} \int_{-m}^{m} \psi(it) \int_{u_1}^{u_2} e^{-itu}\, du\, dt$$

$$= \lim_{m\to\infty} \int_{u_1}^{u_2} \frac{du}{2\pi} \int_{-m}^{m} e^{-itu}\, \psi(it)\, dt$$

$$= \lim_{m\to\infty} \int_{u_1}^{u_2} h_m(u)\, du$$

since $h_m(u)$ converges uniformly to $h(u)$ the limit may be taken inside the integral and

$$H(u_2) - H(u_1) = \int_{u_1}^{u_2} h(u)\, du$$

Since $H(u)$ is continuous at zero and we have shown absolute continuity elsewhere, it is absolutely continuous on an interval including the point zero.

In case $E(Z_1^K) < \infty$ and $r < \gamma + K - 1$, integration by parts of (A) and corollary b to theorem 10 shows that the first r derivatives of $h(u)$ are continuous if $u \neq 0$. The integral expression for $h(u)$ in terms of $\psi(it)$ shows that $\gamma > r + 1$ implies $h^{(r)}(u)$ is continuous at zero. This implies that $G(u) = a + \int_0^u g(v)\,dv$ for $u > 0$; $g(v)$ is continuous for $v \neq 0$. If $E(Z_1^K) < \infty$, then $g^{(r)}(u)$ is continuous for $u \neq 0$ provided $r < \gamma + K - 1$ and is continuous for $u = 0$ provided $r < \gamma - 1$.

Corollary to theorem 11. $p_{nr} = P(z_n = r) \to 0$ uniformly in r, $r \geq 1$ as $n \to \infty$.

Proof: $$p_{nr} = G_n\left(\frac{r}{\mu^n}\right) - G_n\left(\frac{r-1}{\mu^n}\right)$$

$$= \left[G_n\left(\frac{r}{\mu^n}\right) - G\left(\frac{r}{\mu^n}\right)\right] + \left[G\left(\frac{r}{\mu^n}\right) - G\left(\frac{r}{\mu^n} - \frac{1}{\mu^n}\right)\right]$$
$$+ \left[G\left(\frac{r-1}{\mu^n}\right) - G_n\left(\frac{r-1}{\mu^n}\right)\right]$$

Since $G_n(u) \to G(u)$ uniformly for $u > 0$, $G(u)$ is uniformly continuous for $u > 0$ and $G(u)$ is right continuous at $u = 0$, the corollary is proven.

It is interesting to note that Yaglom [22] treated this subject in a manner similar to Harris. Yaglom defined

$$a_n = 1 - p_{no}$$

$$y_n = a_n \frac{z_n}{\mu^n} = a_n w_n$$

$$H_n(y) = P(y_n \leq y \mid y_n > 0)$$

$$H(y) = \lim_{n \to \infty} H_n(y)$$

$$\psi(it) = \int_{-\infty}^{\infty} e^{ity} d H(y)$$ and proved the following theorem.

Theorem 15. If $\mu > 1$, the variance of z_1 is not zero, and $f''(1) < \infty$ then $$f[\psi(it)(1-a) + a] = (1-a)\,\psi(\mu t) + a$$

In this paper we followed the general line of Harris' argument and defined k(s) such that $$\psi(\mu s) = \frac{f[(1-a)\psi(s) + a] - a}{1-a}.$$

This led us to the conclusion that H(u) was the distribution function for $(1-a)w$ which is the limit of $(1-p_{no})\frac{z_n}{\mu^n}$.

Theorem 17. If $\mu < 1$ and $f'' < \infty$ then

$$\lim_{n\to\infty} P^*_{nr} = P^*_r \qquad \sum_{r=1}^{\infty} P^*_r = 1$$

$$\lim_{n\to\infty} F_n(s) = F(s)$$

and $$F[f(s)] = \mu F(s) + (1-\mu) \qquad |s| \le 1$$

$$F(1) = 1 \qquad F'(1) = K^{-1}$$

Proof:

$$F_n(s) = \sum_{r=1}^{\infty} P(Z_n = r \mid Z_n \neq 0)\, s^r$$

$$= \sum_{r=1}^{\infty} \frac{P(Z_n = r, Z_n \neq 0)}{P(Z_n \neq 0)} s^r$$ by definition of conditional probability

$$= \sum_{r=1}^{\infty} \frac{P(Z_n = r)}{P(Z_n \neq 0)} s^r$$ since the summation is over r 0

$$= \sum_{r=1}^{\infty} \frac{P(Z_n = r)}{1 - f_n(0)} s^r$$

$$= \frac{f_n(s) - f_n(0)}{1 - f_n(0)}$$

(1) $$F_n(s) = 1 + \frac{f_n(s) - 1}{1 - f_n(0)}$$

The following theorem may be found in Konig [14]. If $\theta(s)$ is analytic in a domain, R, such that $\theta(s_0) = s_0$ and $|\theta'(s_0)| = a < 1$ for some s_0 in R, then, in this domain $\dfrac{\theta(s) - s_0}{a^n}$ converges uniformly to a continuous function D(s) such that $D[\theta(s)] = a D(s)$

$$D(s_0) = 0 \qquad D'(s_0) = 1$$

$f(s)$ is analytic in $|s| < 1$. Since $f(1) = 1$ and $f'(1) < 1$, $(\mu < 1)$, the s_0 of Konig's theorem is one. We can conclude that $\frac{f_n(s)-1}{\mu^n}$ converges uniformly to a continuous function $D(s)$ such that

(2) $D(f(s)) = \mu D(s)$, $D(1) = 0$, $D'(1) = 1$.

From (1), $F_n(s) = 1 + \left(\frac{f_n(s)-1}{\mu^n}\right)\left(\frac{1}{\frac{1-f_n(0)}{\mu^n}}\right)$

$$\lim_{n\to\infty} F_n(s) = F(s) = \lim_{n\to\infty}\left(1 + \left(\frac{f_n(s)-1}{\mu^n}\right)\left(\frac{1}{\frac{1-f_n(0)}{\mu^n}}\right)\right)$$

(3) $F(s) = 1 + D(s)\,K^{-1}$ by Konig's theorem, since convergence of $\frac{f_n(s)-1}{\mu^n}$ is uniform, the limit of $\frac{1-f_n(0)}{\mu^n}$ exists.

From (3), then, $D(s) = (F(s)-1)K$

From (2), $D(f(s)) = \mu D(s)$

$$= \mu K[F(s)-1]$$

$$F[f(s)] = 1 + \frac{D[f(s)]}{K} \qquad \text{by (3)}$$

$$= 1 + \frac{\mu K[F(s)-1]}{K}$$

$$= 1 + \mu F(s) - \mu$$

And, finally, $F[f(s)] = \mu F(s) + (1-\mu)$

From (3), $F(1) = 1 + D(1)K^{-1}$

$= 1$ since $D(1) = 0$ by (2)

$$F'(1) = \frac{d}{ds}F(s)\Big|_{s=1} = \frac{d}{ds}\left(1 + K^{-1}D(s)\right)\Big|_{s=1} = K^{-1}D'(s)\Big|_{s=1} = K^{-1} \quad \text{by (2)}$$

One of the more interesting questions in the subject of branching processes is the determination of the number of generations to extinction when $\mu < 1$.

Let $N = \min[n : Z_{n+1} = 0]$

$$\theta(s) = \sum_{n=0}^{\infty} P(N=n) e^{ns} = \sum_{n=0}^{\infty} (P_{n+1,0} - P_{n0}) e^{ns}$$

$$b_n = 1 - P_{n+1,0} \qquad b_0 = 1 - P_0$$

$$\theta_1 = \sum_{n=0}^{\infty} b_n e^{ns}$$

<u>Theorem 18.</u> The probabilities b_n satisfy the recursion formula,

$$b_{n+1} = 1 - f(1-b_n)$$

<u>Proof:</u> $b_{n+1} = 1 - P_{n+2,0} = 1 - \sum_{j=0}^{\infty} P_j (P_{n+1,0} S^0)^j$

$$= 1 - \sum_{j=0}^{\infty} P_j (1-b_n)^j$$

$$= 1 - f(1-b_n)$$

<u>Theorem 19.</u> The moment generating function of N is given by

$$\theta(s) = 1 + (e^s - 1)\theta_1(s)$$

<u>Proof:</u> $\theta(s) = \sum_{n=0}^{\infty} (P_{n-1,0} - P_{n0}) e^{ns}$

$$= \sum_{n=0}^{\infty} \left(e^{s(n+1)} - e^{s(n+1)}\right) + \sum_{n=0}^{\infty} \left(P_{n+1,0} e^{ns} - P_{n0} e^{ns}\right)$$

$$= 1 + \sum_{n=0}^{\infty} \left(e^{s(n+1)} - e^{ns} + P_{n+1,0} e^{ns} - P_{n+1,0} e^{(n+1)s}\right)$$

$$= 1 + \sum_{n=0}^{\infty} (e^s - 1) e^{ns} (1 - P_{n+1,0})$$

$$= 1 + (e^s - 1) \sum_{n=0}^{\infty} b_n e^{ns}$$

$$= 1 + (e^s - 1)\theta_1(s)$$

So $\Theta(s)$ is determined once $\Theta_1(s)$ is known. If we let $z = e^s$, $\Theta_1(s)$ is a power series whose coefficients are successive iterates of $f^*(b) = 1 - f(1-b)$ since $b_{n+1} = f^*(b_n) = f^*_{n+1}(b_0)$, where $f^*(0) = 0$ $f^{*\prime}(0) = \mu < 1$. (see Fatou [8].) A function of this sort is meromorphic with poles at $s = -n \log \mu$, $n = 1, 2, \ldots$ So

$$\Theta_1(s) = \frac{X_1 Y_2}{1-\mu e^s} + \frac{X_2 Y_0^2}{1-\mu^2 e^s} + \frac{X_3 Y_0^3}{1-\mu^3 e^s}$$

converges everywhere except at the poles. (see Lattes [29])

$X(s) = \sum_{i=1}^{\infty} x_i s^i$ is determined by $X(\mu s) = f^*(X(s))$, $X''(1) = X'(1) = 1$. y_0 is found by $X(y_0) = b_0 = 1 - p_0$ or by using inverse functions $X^{-1}(f^*(s)) = \mu X^{-1}(s)$ and determining $X^{-1}(b_0)$.

If our process is of the type in which the "parent" does not die, then the total number of individuals is of interest. If $\mu \leq 1$, the probability that this number is finite is equal to one by theorem 4a.

Let $S = \sum_{n=0}^{\infty} Z_n$; then $P(S < \infty) = 1$ for $\mu \leq 1$

Let $q_r = P(S = r)$, $g(s) = \sum_{r=0}^{\infty} q_r s^r$ for $|s| < 1$.

<u>Theorem 20</u>. If $\mu \leq 1$, then $g(s) = s\, f(g(s))$.

<u>Proof</u>: Let $S_n = \sum_{i=0}^{n} Z_n$ $Z_n = 0$ for some n with probability one when $\mu \leq 1$ by theorem 4a. Therefore, S_n converges to some random variable S with probability one. The random variable, S_n, represents the total number of individuals in the generations from the zeroth to the n^{th} generation. Let $S_n = 1 + Y_n$ where $Y_k = \sum_{i=1}^{k} Z_i$, since $Z_0 = 1$.

Let $G_k(S) = \sum_{n=0}^{\infty} q_{kn} S^n$, where $q_{k0} = p_0$ and $q_{kr} = P(Y_k = r)$.

The event $y_k = n$ can be divided into n mutually exclusive events with probabilities:

$P(z_1 = i) P(n-i$ individuals in the following k-1 generations)

for $i = 1, 2, \ldots$

So $\Theta(s)$ is determined once $\Theta_1(s)$ is known. If we let $z = e^s$, $\Theta_1(s)$ is a power series whose coefficients are successive iterates of $f^*(b) = 1 - f(1-b)$ since $b_{n+1} = f^*(b_n) = f^*_{n+1}(b_0)$, where $f^*(0) = 0$ $f^{*\prime}(0) = \mu < 1$. (see Fatou [8].) A function of this sort is meromorphic with poles at $s = -n \log \mu$, $n = 1,2,\ldots$ So

$$\Theta_1(s) = \frac{x_1 y_0}{1-\mu e^s} + \frac{x_2 y_0^2}{1-\mu^2 e^s} + \frac{x_3 y_0^3}{1-\mu^3 e^s}$$

converges everywhere except at the poles. (see Lattes [29])

$X(s) = \sum_{i=1}^{\infty} x_i s^i$ is determined by $X(\mu s) = f^*(X(s)), X''(1) = X'(1) = 1.$

y_0 is found by $X(y_0) = b_0 = 1 - p_0$ or by using inverse functions $X^{-1}(f^*(s)) = \mu X^{-1}(s)$ and determining $X^{-1}(b_0)$.

If our process is of the type in which the "parent" does not die, then the total number of individuals is of interest. If $\mu \leq 1$, the probability that this number is finite is equal to one by theorem 4a.

Let $S = \sum_{n=0}^{\infty} Z_n$; then $P(S<\infty) = 1$ for $\mu \leq 1$

Let $q_r = P(S=r)$, $g(s) = \sum_{r=0}^{\infty} q_r s^r$ for $|s| < 1$.

Theorem 20. If $\mu \leq 1$, then $g(s) = s\, f(g(s))$.

Proof: Let $S_n = \sum_{i=0}^{n} Z_i$ $Z_n = 0$ for some n with probability one when $\mu \leq 1$ by theorem 4a. Therefore, S_n converges to some random variable S with probability one. The random variable, S_n, represents the total number of individuals in the generations from the zeroth to the nth generation. Let $S_n = 1 + Y_n$ where $Y_n = \sum_{i=1}^{n} Z_i$, since $Z_0 = 1$.

Let $G_k(s) = \sum_{n=0}^{\infty} q_{kn} s^n$, where $q_{k0} = p_0$ and $q_{kr} = P(Y_k = r)$.

The event $y_k = n$ can be divided into n mutually exclusive events with probabilities:

$P(z_1 = i)P(n-i$ individuals in the following $k-1$ generations)

for $i = 1,2,\ldots$

Starting with one individual at time zero, the probability of producing n individuals in the next k generations is the coefficient of s^n in $G_k(s)$.

Therefore, $q_{k,n} = \sum_{i=0}^{n} P(y_1 = i)P(n-i \text{ in the next } k-1 \text{ generations})$

$$= \sum_{i=0}^{n} p_i(\text{coefficient of } s^{n-i} \text{ in } (G_{k-1}(s))^i)$$

$$= \sum_{i=0}^{n} p_i(\text{coefficient of } s^n \text{ in } (sG_{k-1}(s))^i)$$

$$= \text{coefficient of } s^n \text{ in } (\sum_{i=0}^{n} p_i(sG_{k-1}(s))^i)$$

$$= \text{coefficient of } s^n \text{ in } (\sum_{i=0}^{\infty} p_i(sG_{k-1}(s))^i)$$

$$= \text{coefficient of } s^n \text{ in } f(sG_{k-1}(s))$$

Therefore, $G_k(s) = \sum_{n=0}^{\infty} q_{k,n}s^n = f(sG_{k-1}(s))$. Clearly, $g_n(s) = sG_n(s)$ since $P(y_n = n) = P(S_n = n+1)$ and $g_n(s) = sf(g_{n-1}(s))$. Taking the limit as n approaches infinity, we see that $g(s) = sf(g(s))$.

Using theorem 20, the expected total number of members of a family is: $g'(1) = \frac{d}{ds}\, s f[g(s)]\Big|_{s=1} = f[g(s)] + s f'[g(s)]g'(s)\Big|_{s=1}$

$$= 1 + f'(1)\, g'(1)$$

Then, $g'(1) = \frac{1}{1-\mu}$

Similarly, $g''(1) = \frac{f''(1) + \mu(1-\mu)}{(1-\mu)^3}$ and the variance of S is:

$$\sigma_S^2 = g''(1) + g'(1) - [g'(1)]^2$$

$$= \frac{f''(1) + \mu(1-\mu)}{(1-\mu)^3} + \frac{(1-\mu)^2}{(1-\mu)^3} - \frac{1-\mu}{(1-\mu)^3}$$

$$\sigma_s^2 = \frac{f''(1) + (1-\mu)(\mu-1) + (1-\mu)^2}{(1-\mu)^3}$$

$$= \frac{f''(1)}{(1-\mu)^3}$$

$$\sigma_s^2 = \frac{\sigma^2 - \mu + \mu^2}{(1-\mu)^3}$$

ESTIMATION

In the introduction we mentioned that the parameters, p_r, must be empiraca.ly determined for the given amount and arrangement of atomic material. The method of maximum likelihood can be applied; however, it seems necessary to assume that certain random variables are observable.

Let z_{mr} be the number of individuals in the m^{th} generation who have exactly r descendents in the $(m+1)^{st}$ generation. Then

$$z_m = \sum_{r=0}^{\infty} z_{mr} = \sum_{r=0}^{\infty} r\, z_{m-1,r} \qquad \text{Let } S_n = \sum_{r=0}^{n} z_r$$

<u>Theorem 21.</u> The maximum likelihood estimates of p_r and μ based on observed values of z_{mr} for $m \leq n$ ($n+1$ observed generations) are

$$\hat{p}_r = \sum_{m=0}^{n} \frac{z_{mr}}{S_n} \qquad \hat{\mu} = \frac{S_{n+1} - 1}{S_n}$$

<u>Proof:</u> In order to construct the likelihood function, first we determine the conditional distribution of z_{mr}, $r = 0,1, \ldots$ given z_m. From our initial assumptions, each member of the m^{th} generation has the same probability distribution associated with his progeny. Therefore the probability of k_o members having no descendents is

$p_0^{k_0} \binom{z_m}{k_0}$ where $\binom{z_m}{k_0}$ is the binomial coefficient and is equal to

$\frac{z_m!}{(z_m - k_0)!\,k_0!}$. Of the remaining $z_m - k_o$ members, say $z_m - k_1$ have one descendent. This probability must be $\binom{z_m - k_1}{k_1} p_1^{k}$. Continuing in this manner we get the multinomial functions:

$$P(z_{mo}, z_{m1}, \ldots, z_{mi}, \ldots \mid z_{m-1,0}, z_{m-1,1}, \ldots, z_{m-1,i}, \ldots)$$

$$= \frac{z_m! \prod_{r=0}^{\infty} p_r^{z_{m,r}}}{\prod_{r=0}^{\infty} (z_{mr})!}$$

Therefore $P(z_{n0}, z_{n1}, \ldots \mid z_{n-1,0}, z_{n-1,1}, \ldots)P(z_{n-1,0}, z_{n-1,1}$

$$\ldots \mid z_{n-2,0}, z_{n-2,1}, \ldots) \ldots P(z_{10}, z_{11}, \ldots z_{01})$$

$$= \prod_{m=0}^{n} \frac{Z_m! \prod_{r=0}^{\infty} p_r^{Z_{m-r}}}{\prod_{r=0}^{\infty} (Z_{mr})!}$$

Due to the Markov properties of the branching process this is the joint distribution of the z_{mr} , $m = 0,1, \ldots$, $r = 0,1, \ldots$

The only factor of the likelihood function which depends on the p_r is

$$\prod_{r=0}^{\infty} P_r^{\sum_{m=0}^{n} Z_{mr}}$$

the logarithm of this expression is $\sum_{r=0}^{\infty} \left(\sum_{m=0}^{n} Z_{mr} \log_e (P_r) \right)$

To maximize this expression subject to the restraint, $\sum_{r=0}^{\infty} P_r = 1$ the method of Lagrangian multipliers may be used.

Let $L = \sum_{r=0}^{\infty} \sum_{m=0}^{n} (Z_{mr} \log P_r - \lambda P_r)$

Then $\frac{\partial L}{\partial P_r} = \sum_{m=0}^{n} \left(Z_{mr}\left(\frac{1}{P_r}\right) - \lambda \right)$

Setting $\frac{\partial L}{\partial P_r} = 0$, we get $\sum_{m=0}^{n} Z_{mr} = \lambda P_r$ for $r = 0,1, \ldots$

But $\sum_{r=0}^{\infty} P_r = 0$; summing the above expression over r gives:

$$\sum_{r=0}^{\infty} \sum_{m=0}^{n} Z_{mr} = \lambda$$

The double sum of the z_{mr} is merely S_n so $\lambda = S_n$ and

$$\hat{P}_r = \frac{\sum_{m=0}^{n} z_{mr}}{S_n}$$

To find $\hat{\mu}$ recall that $\mu = \sum_{r=0}^{\infty} r P_r$ and that $\sum_{r=0}^{\infty} r Z_{mr} = Z_{m+1}$.

Using these facts, it follows that:

$$\hat{\mu} = \frac{\sum_{r=0}^{\infty} r \sum_{m=0}^{n} z_{mr}}{S_n}$$

$$= \frac{\sum_{m=0}^{n} \sum_{r=0}^{\infty} r\, z_{mr}}{S_n}$$

$$= \frac{\sum_{m=0}^{n} Z_{m+1}}{S_n}$$

$$\hat{\mu} = \frac{S_{n+1} - 1}{S_n}$$

ION FLUCTUATION

The number of ion pairs produced by a fast primary particle passing through an absorbing medium is a branching stochastic process (Moyal [18]). In this process the primary particle, which may be a gamma ray, has a certain charge, mass, and velocity or a given frequency; the thickness and atomic properties of the absorbing medium are known. The primary interacts with the atoms in the absorber, freeing electrons by the Compton effect. These electrons are the first generation and each may ionize another atom or atoms depending on their energy; secondary electrons may ionize more atoms, etc. The process ends when all the freed electrons are too slow to produce further ionization. This results in some number of free electrons and, thereby, an equal number of ion pairs.

In determining the distribution of this number we shall make the following assumptions:

1. Successive ionizing collisions are statistically independent.

2. The total energy loss of the primary is very much smaller than its initial energy and, hence, the loss of primary energy may be neglected. Therefore, we are not treating slow primaries or very thick absorbers. We also assume that energy losses due to radiation, nuclear interactions, etc. are negligible.

3. The absorbing medium is homogeneous.

4. The recording of data is delayed sufficiently to allow termination of the process. Photoelectric effects in the gas or chamber walls is neglected.

Unlike the branching processes treated in the previous sections

the probability distribution of the number of progeny per individual in a given generation is dependent upon the generation. Thus one may expect different methods to be used. Moyal has applied the Laplace transform method to solve the Kolmogorov equations of a Markov process and this is outlined below.

Let t be the thickness of the absorber

N = the number of absorber atoms per unit volume

$\sigma(E)$ = total cross section

q = primary ionizing collision rate

q_n = P(n ion pairs are produced in a given collision)

$p_n(t)$ = P(n ion pairs produced in a thickness t)

$M(s,t) = \sum_{n=0}^{\infty} p_n(t)e^{-ns}$, the generating function of $p_n(t)$

Then $(N\,\sigma(E)\,(dE)\,dt)$ is the probability of an energy loss between E and $E + dE$ in an absorber of thickness dt. The primary ionization rate is $q = N\int_0^{\infty}\sigma(E)\,dE$

The $\sum_{n=0}^{\infty} q_n = 1$ and $q(dt)$ = P(an ionizing collision in thickness dt). Since at least one ion pair is produced in a given collision, $q_o = 0$.

Based upon the above assumptions we have

(A) $$P_n(t_1 + t_2) = \sum_{j=0}^{n} P_j(t_1)\,P_{n-j}(t_2)$$

$P_n(\delta t) = (1 - q(\delta t)\,\delta_{no} + (q(\delta t))q_n + o(\delta t)$ where δ_{no} is the Kronecker delta and $O(\delta t)$ goes to zero faster than δt as $\delta t \to 0$.

By assumption 1 and equation (A),

$$M(s, t_1 + t_2) = M(s, t_1)\,M(s, t_2)$$

$$\frac{\partial M(s,t)}{\partial t} = \lim_{\delta t \to 0} \frac{M(s,t+\delta t) - M(s,t)}{\delta t}$$

$$= \lim_{\delta t \to 0} \frac{1}{\delta t}\left(M(s,\delta t) - 1\right) M(s,t)$$

$$= \lim_{\delta t \to 0} \frac{1}{\delta t}\left[1 - q(\delta t) + q(\delta t)\left(\sum_{n=1}^{\infty} q_n e^{-ns}\right) - 1\right] M(s,t)$$

$$= \lim_{\delta t \to 0} q M(s,t)\left[\sum_{n=1}^{\infty} q_n e^{-ns} - 1\right]$$

$$\frac{\partial M(s,t)}{\partial t} = M(s,t)\left[q \sum_{n=1}^{\infty} q_n (e^{-ns} - 1)\right)$$

Comparing this equation with $\frac{\partial y}{\partial x} = cy$, it is clear that

$$M(s,t) = \exp\left[qt \sum_{n=1}^{\infty} (e^{-ns} - 1) q_n\right] = e^{QR(s)}$$

where $\quad Q = qt \qquad\qquad R(s) = \sum_{n=1}^{\infty} (e^{-ns} - 1) q_n$

The inversion of the Laplace transform, M(s,t), is given by:

$$P_n(Q) = \frac{1}{2\pi i} \int_{c-\pi i}^{c+\pi i} e^{QR(u)} e^{nu} \, du \qquad \text{(see Widder [21])}$$

Evaluating this integral by the saddlepoint method (see Korn [15]), using the first term only, gives:

$$P_n(Q) = \frac{1}{c}\left[2\pi Q R''(s_n)\right]^{-1/2} \exp\left[Q\left(R(s_n) - S_n R'(s_n)\right)\right]$$

where S_n is related to n by

$$n = -QR'(s_n) = Q \sum_{K=1}^{\infty} (K e^{-KS_n}) (q_K)$$

since $$R(S_n) = \sum_{K=1}^{\infty} (e^{-S_n K} - 1) q_K$$

$$R'(S_n) = -\sum_{K=1}^{\infty} K q_K e^{-S_n K}$$

$$R''(S_n) = \sum_{K=1}^{\infty} K^2 q_K e^{-S_n K}$$

and c is a normalizing constant.

Since $$\sum_{n=0}^{\infty} P_n(Q) = 1$$

$$c = \sum_{n=0}^{\infty} \left[2\pi Q R''(S_n)\right]^{-1/2} \exp\left[Q(R(S_n) - S_n R'(S_n))\right]$$

APPENDIX

Bibliography

BIBLIOGRAPHY

1. Arley, Niels, "On the Theory of Stochastic Processes and Their Applications to the Theory of Cosmic Radiation", G.E.C. Gads Forlag, Copenhagen, 1943.

2. Bharucha-Reid, A. T , "Elements of the Theory of Markov Processes and Their Applications", McGraw-Hill Book Co., Inc., New York, N.Y., 1960.

3. Cramer, H., "Random Variables and Probability Distributions", Cambridge Tract 36, 1938.

4. Doob, J. L., "Stochastic Processes", John Wiley and Sons, Inc., New York, N.Y., 1960.

5. Fatou, P., "Sur une classe remorkable de series de Taylor", Annales Sci. de l'Ecole Normale Sup. de Paris, vol. 27, series 3, pp. 43 - 53, 1910.

6. Feller, W., "An Introduction to Probability Theory and Its Applications", vol. 1, Second edition, John Wiley and Sons, Inc., New York, N.Y., 1957.

7. Fisher, R. A., "The Genetical Theory of Natural Selection", Oxford University Press, New York, N.Y.,1930.

8. Galton, F. and H.W. Watson, "On the Probability of the Extinction of Families", Journal of the Anthropological Institute, vol. 4, pp. 138 - 144, 1874.

9. Harris, T. E.,"Some Theorems on the Bernoullian Multiplicative Process", Dissertation, Princeton, 1947.

10. ______, "Branching Processes," Annals of Mathematical Statistics, vol. 19, pp. 474 - 494, 1948.

11. Hawkins, D. and S. Ulam, "Theory of Multiplicative Processes, I, Los Alamos Scientific Laboratory Declassified Document 265, 1944.

12. Kendall, D. G., "Stochastic Processes and Population Growth", Journal of the Royal Statistical Society, ser B, vol. 11, pp 230 - 264, 1949.

13. Kolmogorov, A., "Grundbegriffe der Wahrscheinlichkeitsrechnung, Chelsea, 1946.

14. Konigs, G., "Nouvelles recherches sur les integrals des certaines equations fonctionnelles," Ann. sci. ecole norm. super. de Paris, ser. 3, vol. 1, pp. 3 - 41, 1884.

15. Korn, G. A. and T. M. Korn, Mathematical Handbook for Scientists and Engineers, McGraw-Hill Book Co., Inc., New York, N.Y., 1961.

16. Lattes, S. "Sur les suites recurrentes non lineaires et sur les functions generatrices de ces suites", Annales de la Fac. des Sciences de l'Universite de Toulouse, vol. 3, series 3, pp. 96 - 105, 1911.

17. Lotka, A. J., "Theorie analytique des associations biologiques", 2, Herman, Paris, 1930.

18. Moyal, J. E., "Theory of Ionization Fluctuations", The Philosophical Magazine, ser. 7, vol. 46, pt. 1, no. 374, p. 263, 1955.

19. Otter, R., "The Multiplicative Process, Annals of Mathematical Statistics, vol. 20, pp. 206 - 224, 1949.

20. Urbanik, K.,"On a Stochastic Model of a Cascade", Bull. acad. polon. sci., Classe III, vol. 3, pp. 349 - 351, 1955.

21. Widder, D. V., "The Laplace Transform", Princeton University Press, Princeton, N. J., 1941.

22. Yaglom, A. M., "Certain Limit Theorems of the Theory of Branching Stochastic Processes, Doklady Akad. Nauk, S.S.S.R., n.s., vol. 56, pp. 795 - 798, 1947. (See Math Rev., vol. 9, no. 3, p. 149, 1948)